



Avaliação Didática para Concurso de Livre Docência

*Abstrações de Dados e suas
Representações em Modelos de Dados*

João Eduardo Ferreira (jef@ime.usp.br)



IME - USP



Agenda

- Motivação
- Principais Abstrações
- Abstrações e suas Representações
 - Exercícios
- Fronteiras
- Próxima Aula





“Os DEZ MANDAMENTOS”

- I. Adorarás as abstrações de dados sobre todas as coisas;
- II. Não Projetarás bancos de dados baseado no modelo físico;
- III. Projetarás primeiramente as classes “naturais”; (não cairás na tentação de classes complexas originárias de relacionamentos e muito menos confundirás atributos com classes e, ainda, relacionamentos com classes);
- IV. Não confundirás relacionamentos binários com triplos;
- V. Não confundirás relacionamento triplo com agregação;
- VI. Não confundirás composição lógica com física;



“Os DEZ MANDAMENTOS”

- VII. Compreenderás a evolução dos relacionamentos nas hierarquias de especialização e composição;
- VIII. Não utilizarás o mapeamento das abstrações de dados em vão;
(Conhecerás as formas de mapeamento objeto-relacional. Não cairás na tentação de mapeamentos fúteis);
- IX. Não levantarás falso testemunho sobre o valor semântico dos dados;
(Não ignorarás a capacidade do usuário como solucionador de problemas na identificação do valor semântico dos dados);
- X. Rejeitarás a tentação do mundo estático;





Bibliografia

- Elmasri, R.; Navathe, S. B. [Trad.]. Sistemas de bancos de dados. Traduzido do original: FUNDAMENTALS OF DATABASE SYSTEMS. São Paulo: Pearson(Addison Wesley), 2005. 724 p. ISBN: 85-88639-17-3. Parte 1, páginas 1 – 81.
- Heuser, C.A., Projeto de Banco de Dados., Sagra - Luzzatto, 1 edição, 2004. Páginas 1 – 45.
- Teorey, T.J.; Lighstone, S.; Nadeau, T. [Trad.] Projeto e Modelagem de Banco de Dados. Traduzido do original: Database Modeling and Design, 6rd Ed., Morgan Kaufmann, San Francisco, Calif., 2006.
- Conceptual Database Design Batini, Ceri e Navathe Benjamin/Cummings Pub. Co. 1992.
- Takai, O.K; Italiano, I.C.; Ferreira, J.E. Introdução a Banco de Dados. Apostila. Site: <http://www.ime.usp.br/~jef/apostila.pdf>





Pressupostos

- Que os alunos leram os textos mencionados na bibliografia.
- Que os alunos saibam os conceitos de:
 - Banco de Dados
 - Modelo de Dados
 - SGBD
 - Modelagem de Dados
 - Modelagem Conceitual
 - Projeto Lógico e Físico de BD

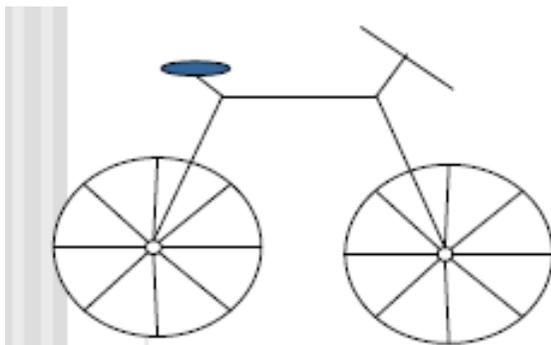




O que é Abstração de Dados ?

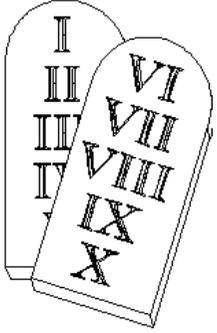
- Processo mental que seleciona algumas propriedades de um conjunto de objetos e exclui outras irrelevantes em um dado contexto.

Def. do Michaelis: Velocípede de duas rodas iguais, movido a pedal .



a figura de uma bicicleta é uma representação dessa abstração ! O texto com uma definição de dicionário também é.





O que são Abstrações de Dados ?

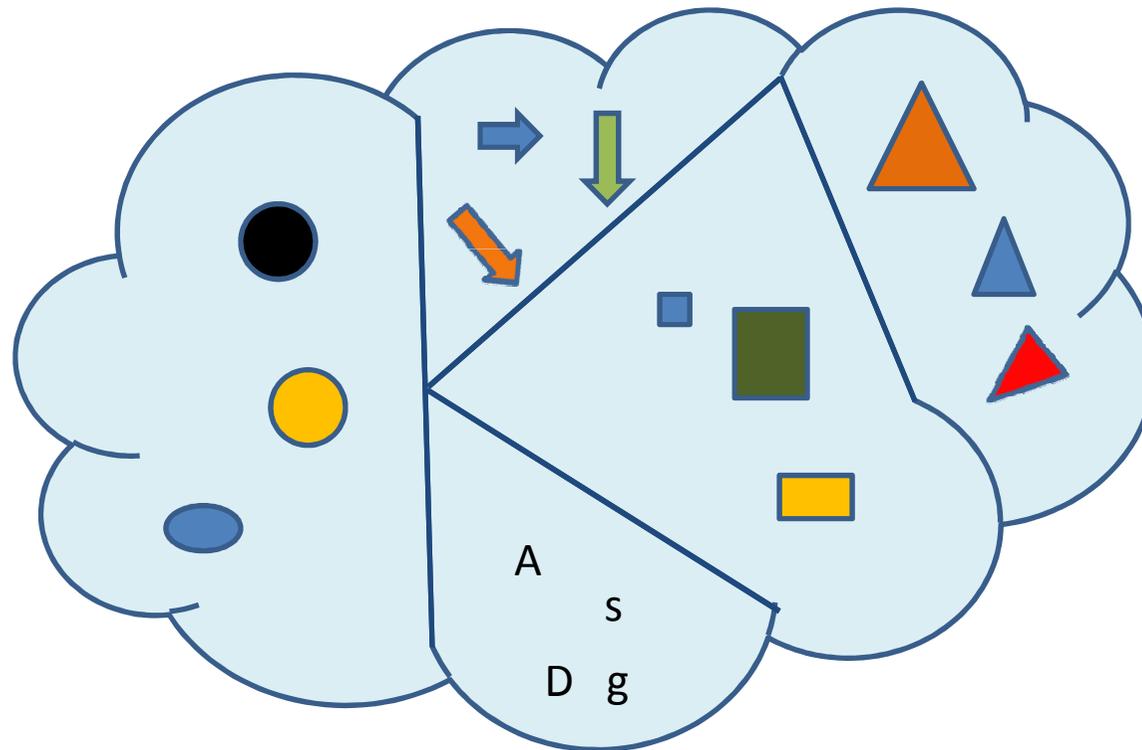
- Proposto inicialmente por Smith e Smith [1977].
- Aplicado ao modelo ER por Scheuermann, Scheffner e Weber [1980], Elmasri e Navathe [2003], Bruce [1992], IDEFIX [2005], entre outros.
- A aplicação do modelo de rede semântica à modelagem conceitual foi mostrada por Bachman [1977], McKleod e King [1979], Hull e King [1987] e Peckham e Maryanski [1988].





Principais Abstrações

Classificação



Particionamento (disjunto) do minimundo de acordo com algum critério de similaridade.

Controle de Transações e Indexação de Banco de Dados

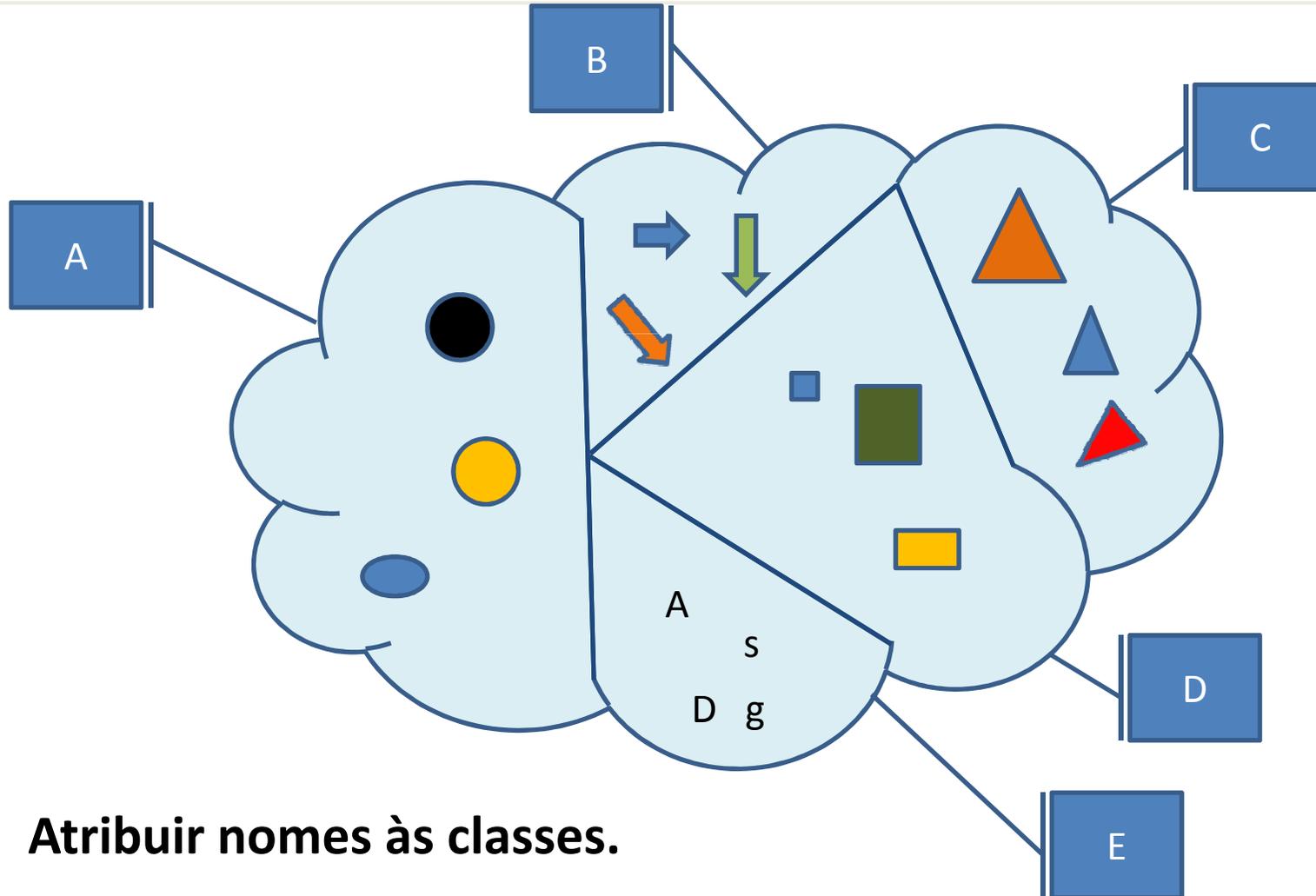


IME - USP



Principais Abstrações

Classificação



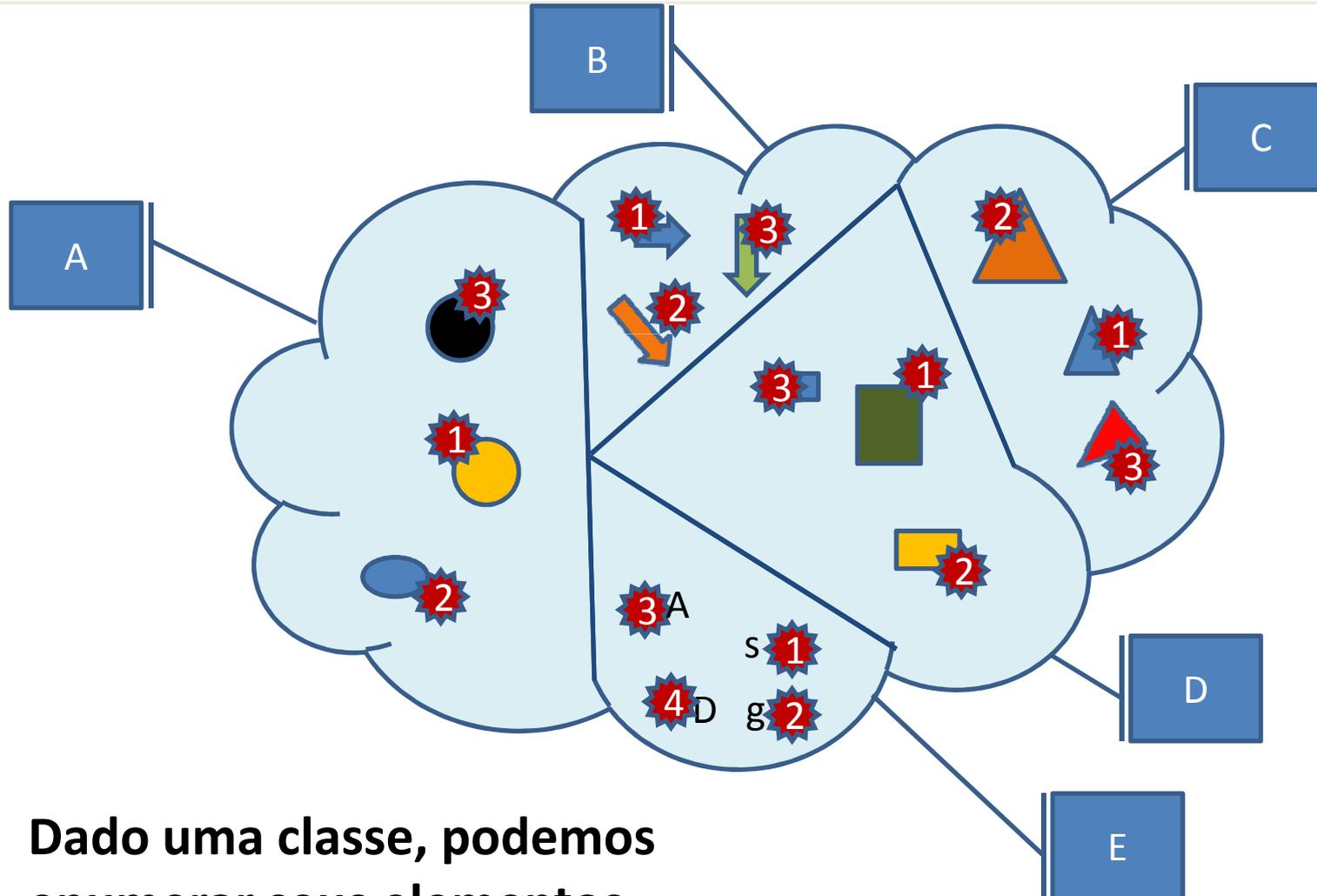
Atribuir nomes às classes.





Principais Abstrações

Classificação



Dado uma classe, podemos enumerar seus elementos.

Controle de Transações e Modelagem de Banco de Dados



IME - USP



Principais Abstrações

Classificação

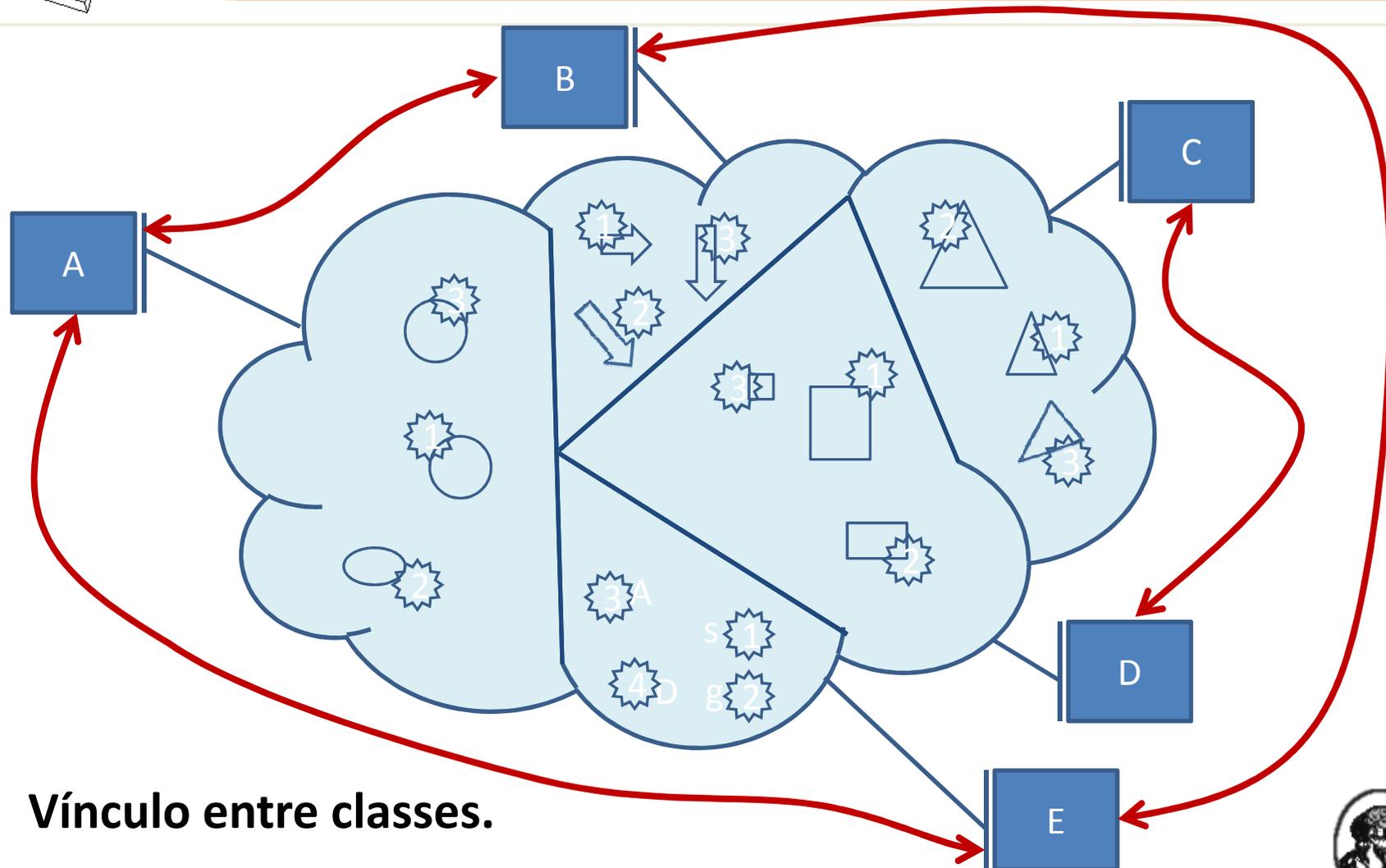
- **Classe:** conjunto de objetos descritos pelo mesmo conjunto de propriedades.
- **Tipo:** conjunto de propriedades que descreve os objetos de uma classe.
- **Instância:** elemento de uma Classe.
- **Abstração por Classificação:**
 - Define um conceito abstrato a partir do levantamento das propriedades comuns a uma coleção de objetos do mundo real.
 - Estabelece um relacionamento do tipo “é membro de” entre a classe e suas instâncias.





Principais Abstrações

Relacionamento



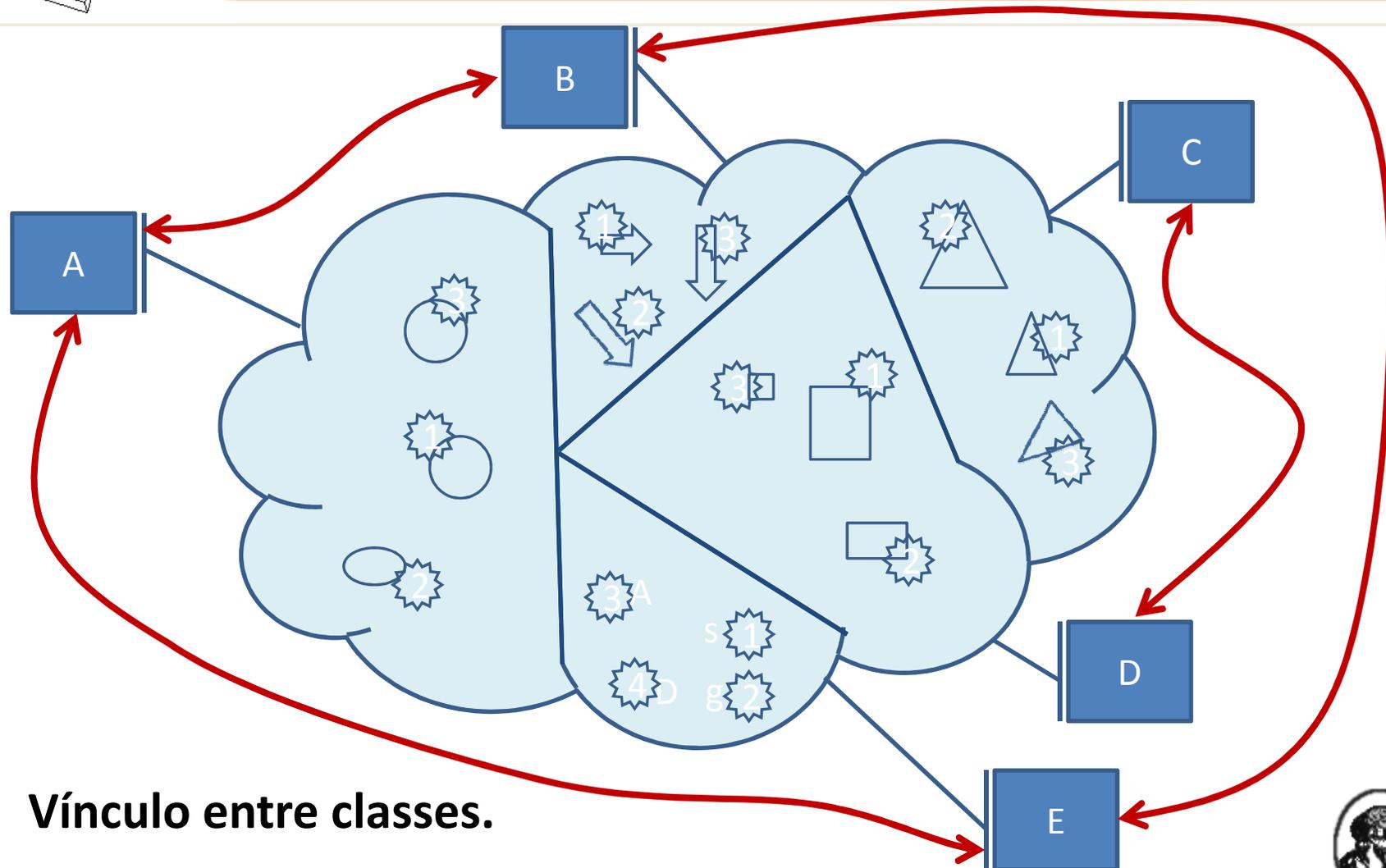
Vínculo entre classes.





Principais Abstrações

Participação e Cardinalidade



Vínculo entre classes.





Principais Abstrações

Vínculos Evoluem

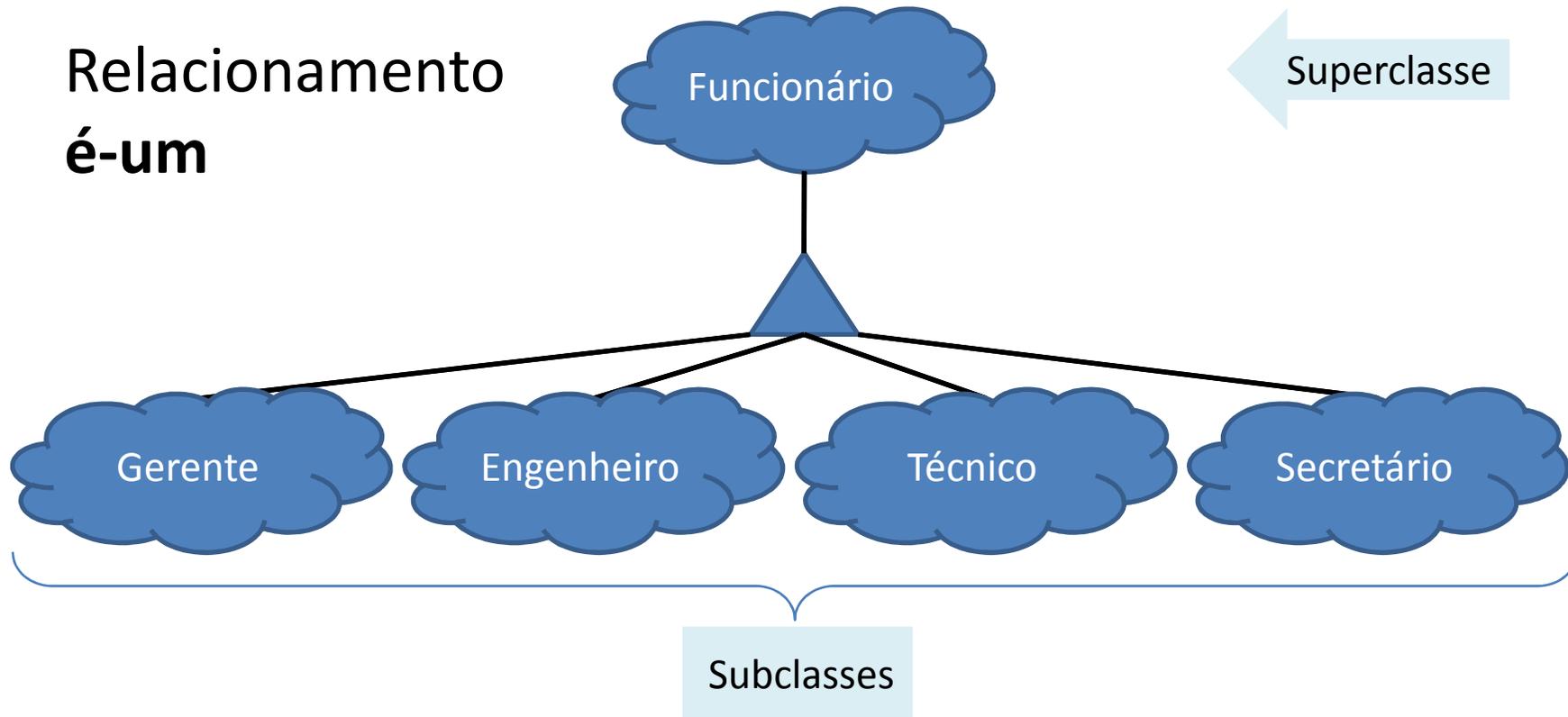
- Especialização
- Composição
- Agregação



Principais Abstrações

Especialização

Relacionamento
é-um





Principais Abstrações

Especialização

- **Subclasse:** É uma Classe cujas instâncias devem sempre ser um subconjunto das instâncias de sua superclasse, configurando um relacionamento denominado Superclasse / Subclasse.
- **Especialização** $Z = (S_1, S_2, \dots, S_n)$: É um conjunto de subclasses que possuem a mesma superclasse G , ou seja, G/S_i é um relacionamento superclasse / subclasse para $i=1, 2, \dots, n$.
- **Superclasse:** É uma Classe que possui relacionamento G/S .





Principais Abstrações

Especialização

- **G** é chamada de classe generalizada, ou uma generalização das subclasses $\{S_1, S_2, \dots, S_n\}$.
- **Z** é considerado total se sempre (em qualquer momento da especialização) temos:
 - $\sum S_i = G$; $i=1, 2, \dots, n$
 - Caso contrário, Z é considerado parcial.
 - Z será disjunto sempre que:
 - $S_i \cap S_j = \emptyset$ para $i \neq j$
 - Caso contrário Z é sobreponível

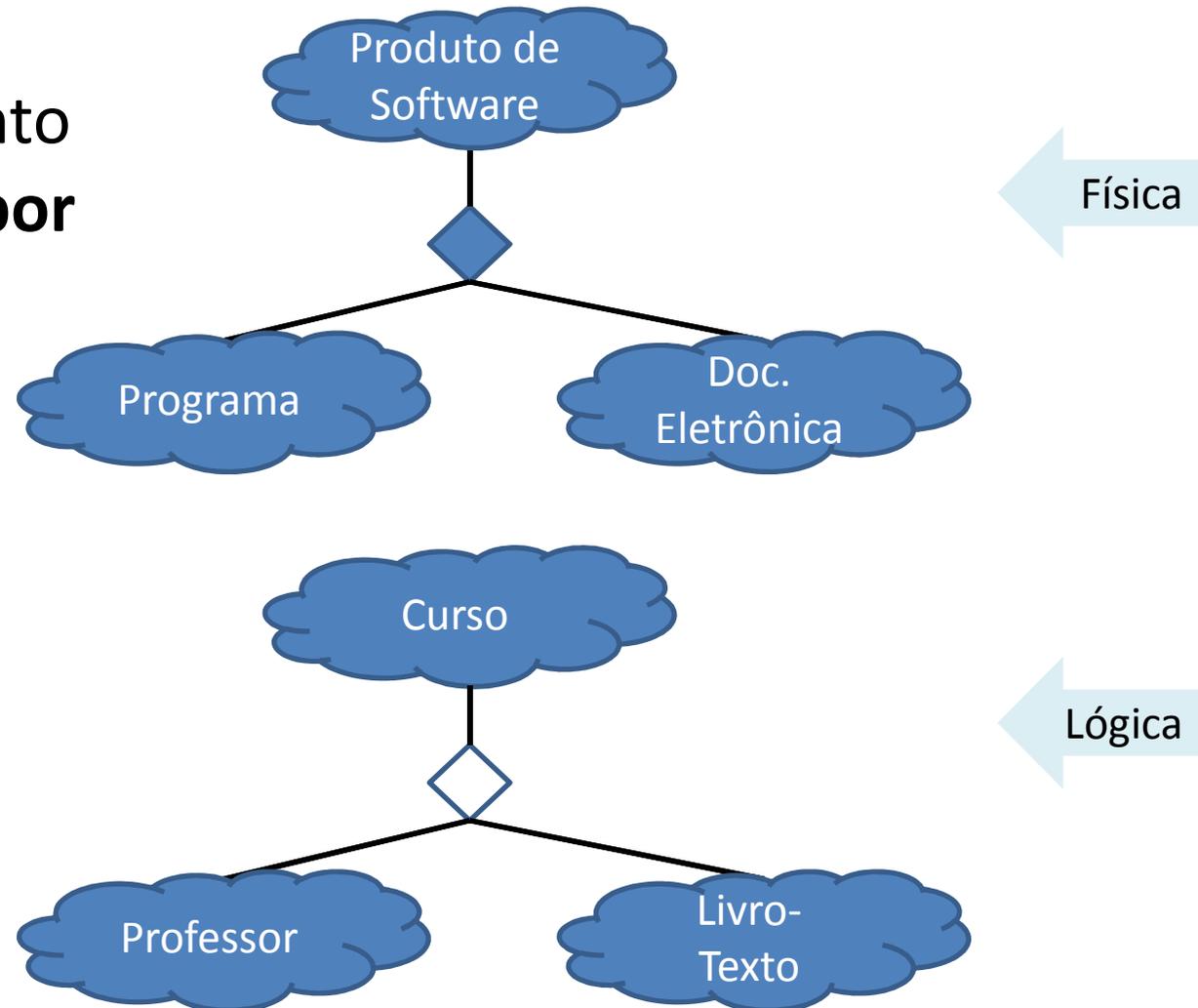




Principais Abstrações

Composição

Relacionamento
é-composto-por





Principais Abstrações

Especialização X Composição

- Composição é uma forma de abstração entre um supertipo e a entidade subtipo e é significativamente diferente da abstração de generalização.
- A generalização é descrita em termos do relacionamento “é um” entre o subtipo e o supertipo. Por exemplo, um Funcionário é um Indivíduo.
- A composição por outro lado, é o relacionamento entre o todo e suas partes é descrita como um relacionamento “parte de”. Por exemplo, um relatório e um protótipo do pacote de software são ambos parte dos produtos de um contrato.

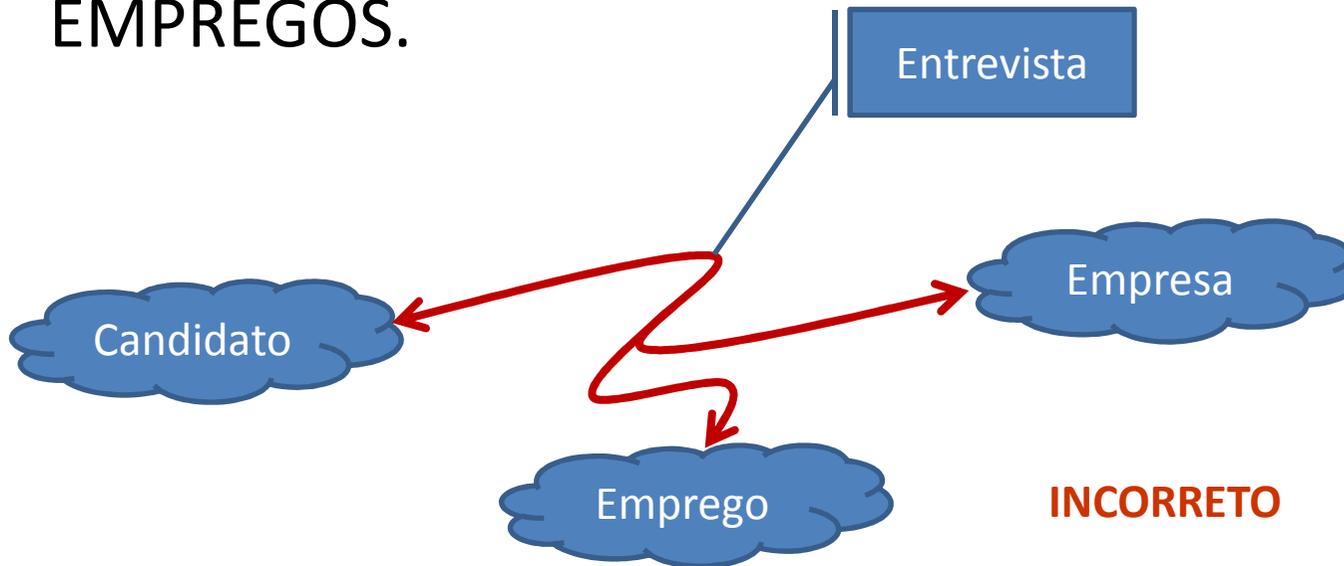




Principais Abstrações

Agregação (Entidade Associativa)

- Considere que algumas entrevistas resultem em EMPREGOS.



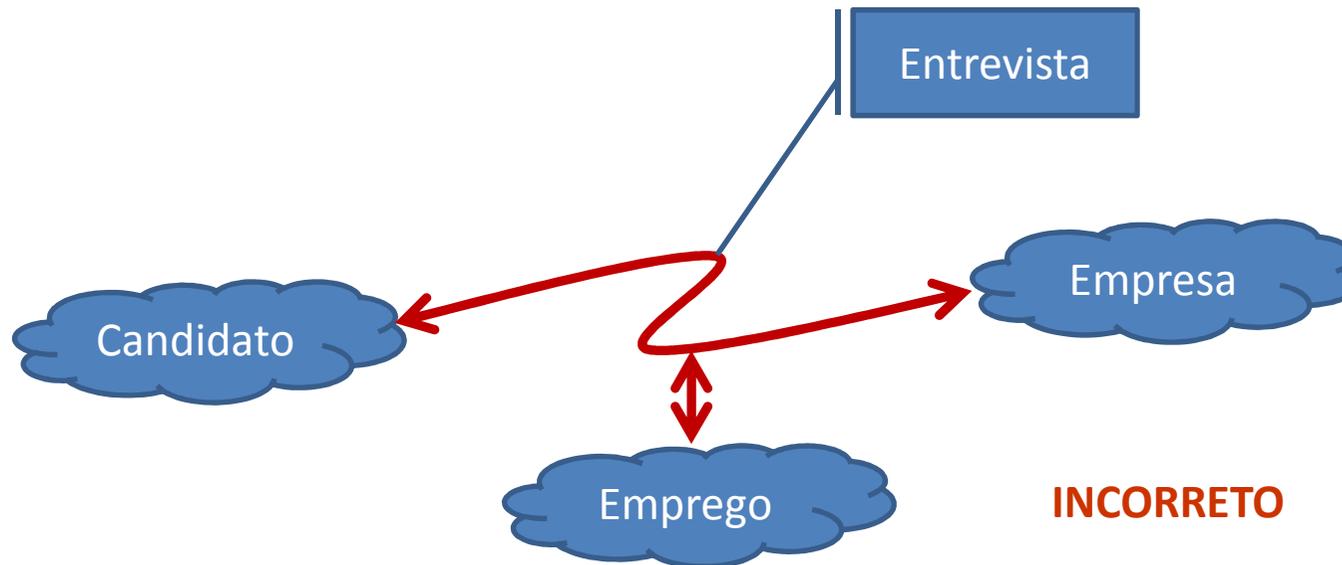
A figura indica que todo relacionamento de ENTREVISTA tem um emprego associado.



Principais Abstrações

Agregação

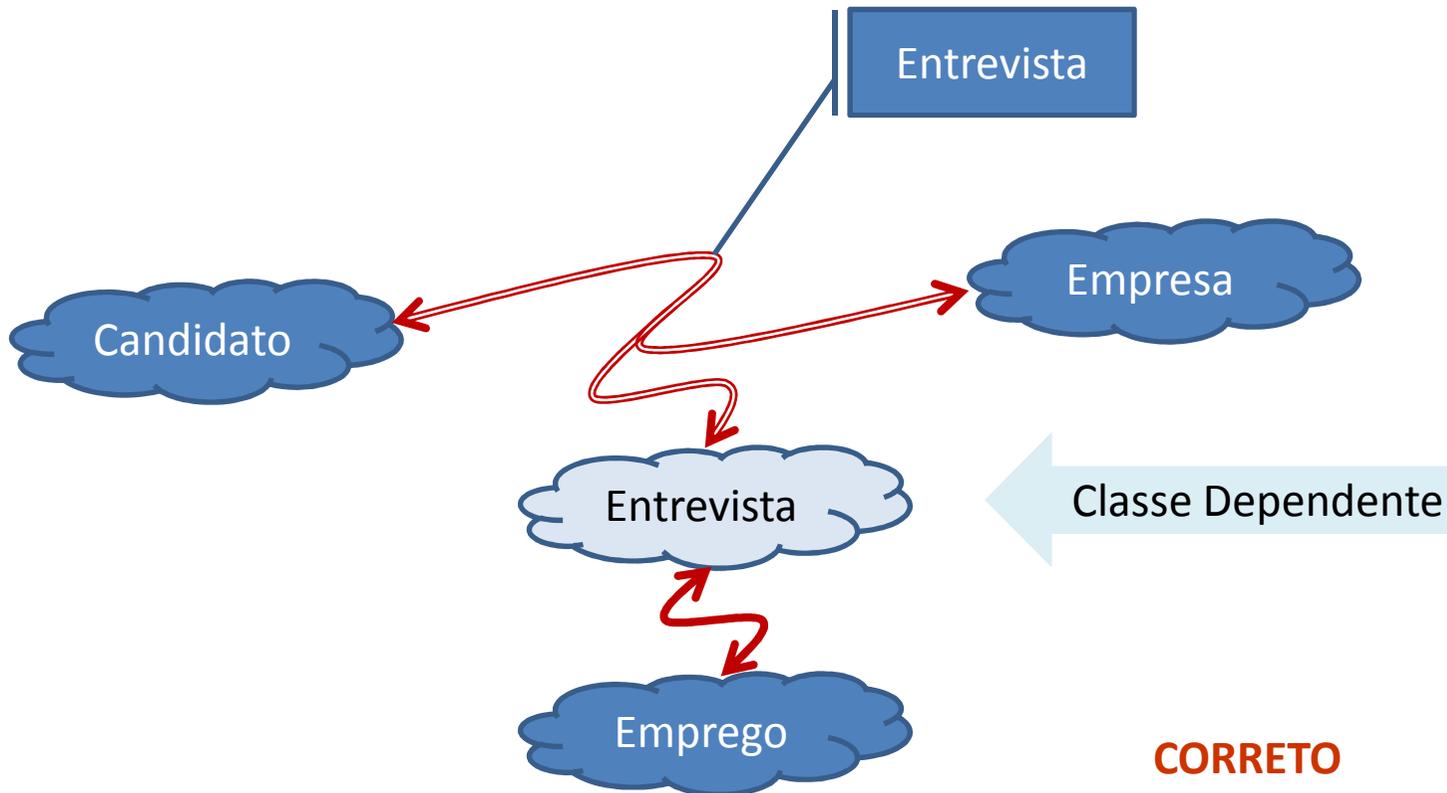
- Também não é correto estabelecer vínculos entre vínculos.





Principais Abstrações

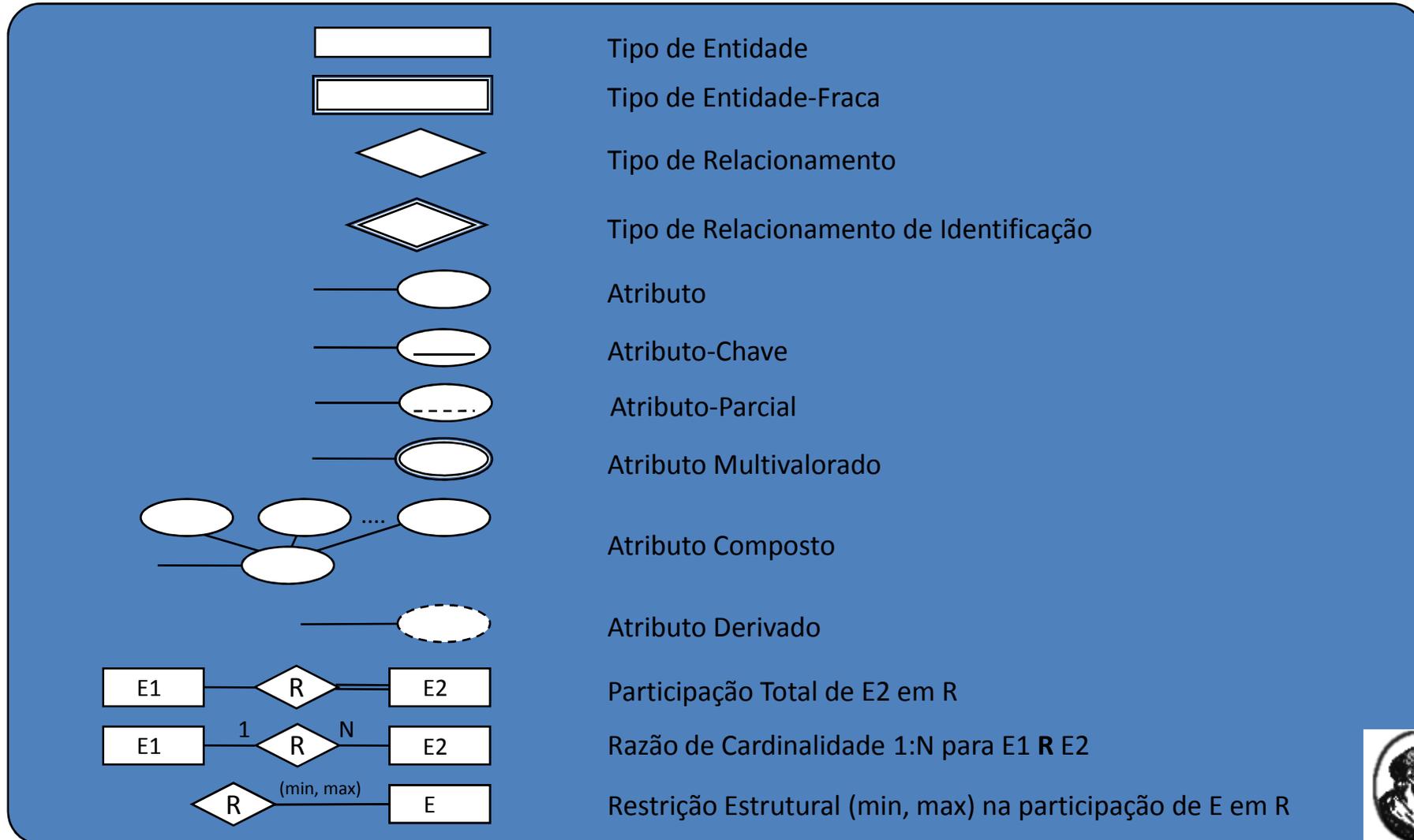
Agregação





Abstrações e suas Representações

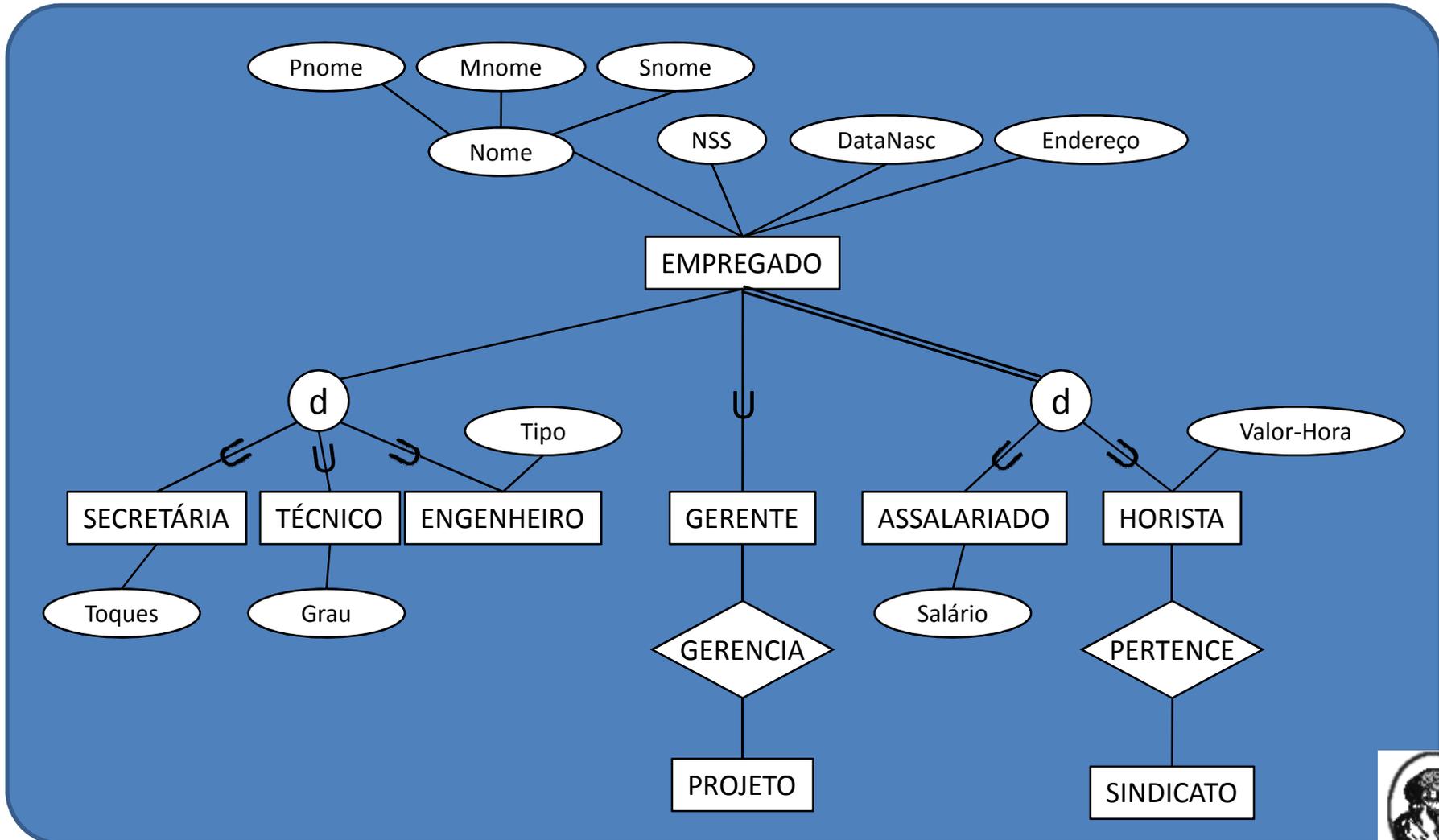
Modelo Entidade-Relacionamento





Abstrações e suas Representações

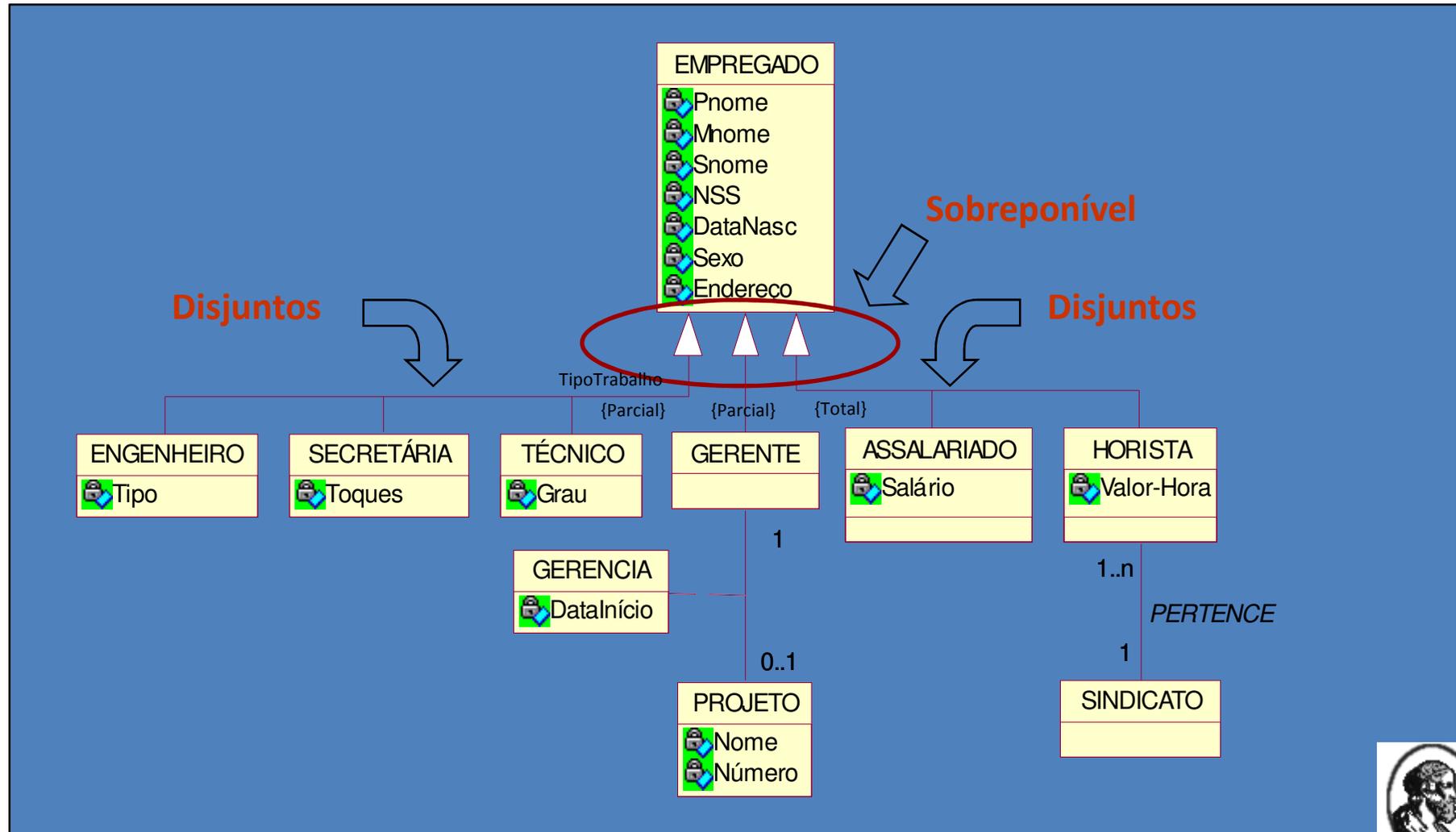
Especialização no Modelo Entidade Relacionamento





Abstrações e suas Representações

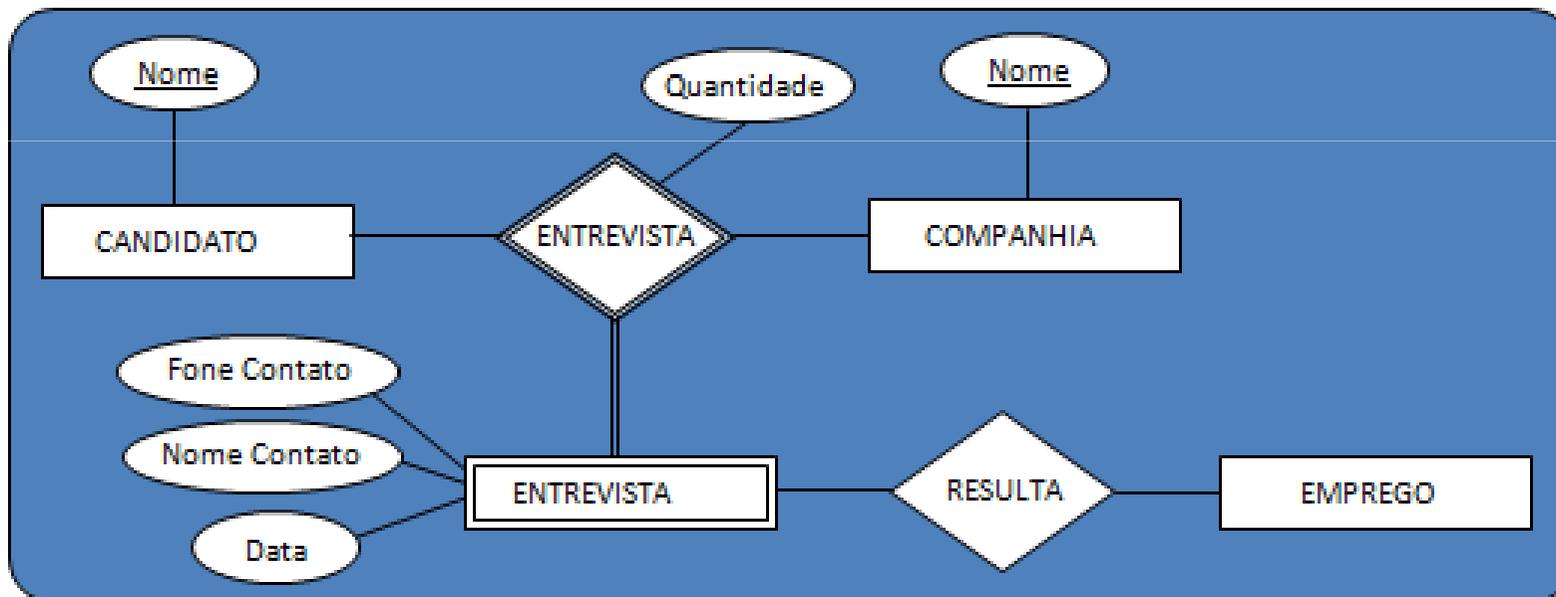
Especialização no Modelo Orientado a Objetos (UML)





Abstrações e suas Representações

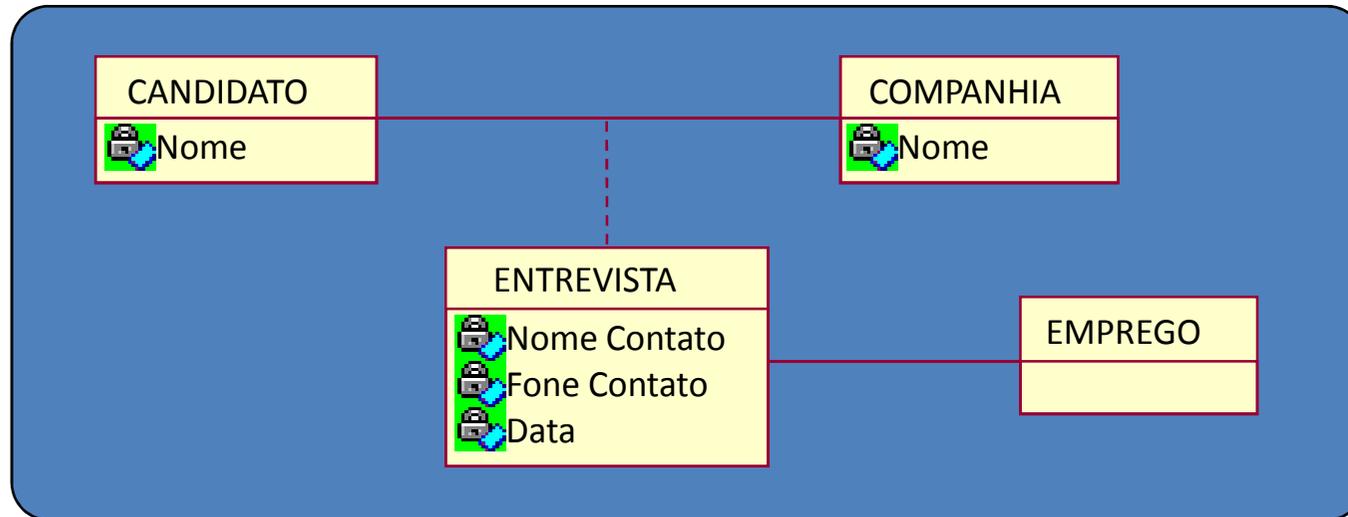
Agregação no Modelo Entidade-Relacionamento





Abstrações e suas Representações

Agregação no Modelo Orientado a Objetos (UML)





Abstrações e suas Representações

Composição no MER e no Modelo Orientado a Objetos (UML)

- No MER-X, não existe uma representação específica para a composição. Esta abstração é representada utilizando relacionamentos e não há com diferenciar entre composição lógica e física. (Obs. Teorey usa “indevidamente” o termo agregação para representar a composição. Página 26).





Abstrações e suas Representações

Composição no MER e no Modelo Orientado a Objetos (UML)

- Na UML existe como representar uma composição (aqui também o termo utilizado é agregação) de duas formas:
 - **Composição-Física** (“Agregação-Forte”): REPRESENTADO por um losango preenchido no lado da classe que compõem objetos de outra classe. Em linguagens OO, classes compostas criam seus componentes dentro da própria classe.
 - **Composição-Lógica** (“Agregação-Fraca”): REPRESENTADO por um losango não-preenchido no lado da classe que compõem objetos de outra classe. Aqui objetos são apenas referenciados pela classe que os compõem.





Abstrações e suas Representações

Mapeamento da Especialização do MER-X para o Modelo de Dados Relacional

- Há várias opções de mapeamento
- A idéia é continuar com a seqüência dos passos de mapeamento DER / MDR
- C denotará a superclasse
- S denotará a subclasse
- $Atr(R)$ denotará os atributos da relação R
- $PK(R)$ denotará a chave-primária de R
- k é chave primária
- $\{k, a_1, a_2, \dots, a_n\}$ são atributos de C.

Fonte: Elmasri & Navathe. Capítulo 7, Página 137.





Abstrações e suas Representações

Mapeamento da Especialização do MER-X para o Modelo de Dados Relacional

- Passo 8a
 - Relações Múltiplas (relações para classes e subclasses).
 - Crie uma relação L para C com:
 - $Atr(L) = \{k, a_1, a_2, \dots, a_n\}$
 - $PK(L) = k$.
 - Crie uma relação L_i para cada subclasse S_i , com:
 - $Atr(L_i) = \{k\} \cup \{\text{atributos de } S_i\}$
 - $PK(L_i) = k$.
- Esta opção funciona para qualquer especialização (total ou parcial, disjuntas ou sobrepostas).
- No caso de disjuntas, há necessidade de busca por k em nas L_i , até encontrá-la.



Abstrações e suas Representações

Mapeamento da Especialização do MER-X para o Modelo de Dados Relacional

- Passo 8b
 - Relações Múltiplas (somente relações para subclasses).
 - Crie uma relação L_i para cada subclasse S_i , com:
 - $Atr(L_i) = \{\text{atributos de } S_i\} \cup \{k, a_1, a_2, \dots, a_n\}$
 - $PK(L_i) = k$.
 - Esta opção funciona somente para especializações cujas subclasses são totais (toda entidade em uma superclasse deve pertencer a, pelo menos, uma subclasse).





Abstrações e suas Representações

Mapeamento da Especialização do MER-X para o Modelo de Dados Relacional

- Passo 8c
 - Relação única com um atributo tipo
 - Crie uma única relação L com:
 - $Atr(L) = \{k, a_1, a_2, \dots, a_n\} \cup \{Atr(S_i)\} \cup \dots \cup \{Atr(S_m)\} \cup \{t\}$
 - $PK(L) = k.$
 - O atributo t é chamado atributo tipo (ou discriminador), que indica a subclasse à qual cada tupla pertence, se pertencer a alguma.
 - Esta opção funciona para especializações cujas subclasses sejam **disjuntas**; mas pode gerar muitos valores nulos se houverem muitos atributos específicos nas subclasses.





Abstrações e suas Representações

Mapeamento da Especialização do MER-X para o Modelo de Dados Relacional

- Passo 8d
 - Relação única com múltiplos atributos de tipo.
 - Crie uma única relação L com:
 - $Atr(L) = \{k, a_1, a_2, \dots, a_n\} \cup \{Atr(S_i)\} \cup \dots \cup \{Atr(S_m)\} \cup \{t_1, t_2, \dots, t_m\}$.
 - $PK(L) = k$.
 - Cada atributo t_i é um atributo do tipo booleano indicando se a tupla pertence ou não à aquela subclasse S_i .
 - Esta opção funciona para especializações cujas subclasses sejam **sobrepostas** (embora também funcione para as especializações disjuntas).



Abstrações e suas Representações

Exercício 1 (Obrigatório)

- Mapeamento da Agregação DER-X para o modelo relacional.





Abstrações e suas Representações

Exercício 2 (Obrigatório)

Abstração + DER-X + Relacional

- Uma federação de futebol pretende desenvolver um sistema informatizado para controlar as informações geradas em seus campeonatos. As informações a serem armazenadas dizem respeito aos times, jogadores e partidas. Um time pode ser profissional ou amador. Para os times amadores deseja-se armazenar nome, cidade de origem e ano em que foram criados. Para os times profissionais deseja-se saber quem são os atuais patrocinadores e nome e sobrenome dos atuais diretores. Para cada jogador deseja-se conhecer em quais times jogou ou joga e, para cada time, em quais posições atuou ou atua. Sobre as partidas de futebol são armazenadas: o placar, o local (cidade), a data, o árbitro principal e, é claro, os times que jogaram. Sobre os árbitros, reconhecidos pela federação, são armazenados nome, RG, o tipo de habilitação que possuem (para apitar partidas regionais, estaduais, nacionais ou internacionais) e, para consulta das torcidas organizadas, a filiação materna 😊. Outra informação importante diz respeito aos uniformes dos times: cada time possui exatamente um uniforme titular e um reserva. Sobre os uniformes também são armazenadas as cores e uma descrição geral.





Abstrações e suas Representações

Exercício 3 (Obrigatório)

Abstração + DER-X + Relacional

- Uma empresa adota o sistema de vendas por meio de representantes de vendas. Cada representante tem uma região de atuação, designada por um nome, à qual estão associadas informações como: número de clientes potenciais e número de clientes já cadastrados. De cada representante de vendas são armazenadas informações como nome, RG, idade e escolaridade. Cada representante atua em apenas uma região, e tem a seu cargo exclusivo os clientes dessa região. Os clientes são estabelecimentos comerciais que pertencem a apenas uma região. De cada cliente, potencial ou cadastrado, são mantidos nome, CNPJ, endereço e contatos já efetuados. Um cliente é cadastrado quando é efetuada a primeira venda para ele. Para cada cliente cadastrado são armazenadas as informações sobre todas as vendas já efetuadas para este cliente. Cada vez que é feito um contato com cliente são registrados a data e o tipo de negócio tratado. Para cada venda efetuada são registrados data, valor, produtos vendidos e quantidade. De cada produto são armazenados um código, descrição e preço corrente.





Abstrações e suas Representações

Exercício Indicados

- Heuser, página 42.
- Elmasri & Navathe, página 56.



Fronteira

+ 6 opções de mapeamento de especialização DER-X para Relacional

- A “estória” do exercício de mapeamento.
<http://www.ime.usp.br/~jef/tentativa-ER2000.pdf>





Fronteira

Um interessante paper

- <http://www.ime.usp.br/~jef/lander.pdf>



Pergamon

Information Systems Vol. 25, Nos. 6–7, pp. 399–415, 2000
© 2000 Published by Elsevier Science Ltd.
Printed in Great Britain
0306-4379/00 \$20.00

PII: S0306-4379(00)00025-9

ON THE RELATIONAL REPRESENTATION OF COMPLEX SPECIALIZATION STRUCTURES[†]

ALTIGRAN S. DA SILVA¹, ALBERTO H.F. LAENDER¹, and MARCO A. CASANOVA²

¹Departamento de Ciência da Computação, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627
31270-010, Belo Horizonte MG, Brasil

²Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rua Marquês de São
Vicente 255, 22453-900, Rio de Janeiro RJ, Brasil

(Received 24 April 1999; in final revised form 16 May 2000)

Abstract — The mapping of entity-relationship schemas (ER schemas) that contain complex specialization structures into the relational model requires the use of specific strategies to avoid inconsistent states in the final relational database. In this paper, we show that for this mapping to be correct it is required to enforce a special kind of integrity constraint, the *key pairing constraint (KPC)*. We present a mapping strategy that use simple inclusion dependencies to enforce KPC and show that this strategy can be used to correctly map specialization structures that are more general than the simple specialization structures considered by previous strategies. © 2000 Published by Elsevier Science Ltd.

Key words: ER Model, Relational Database Design, Specialization Structures, Integrity Constraints



IME - USP



Fronteira

Naked Objects

- O uso do ***naked object*** como valorização do usuário solucionador de problemas na caracterização e delimitação das abstrações de dados.
(<http://www.ime.usp.br/~jef/NOportugues.pdf>)





Fronteira Naked Objects

- <http://www.ime.usp.br/~jef/broinizi-SAC2008.pdf>

Using Annotations in the Naked Objects framework to explore data requirements

Marcos E. B. Broinizi
Department of Computer
Science
University of São Paulo
Rua do Matão, 1010 - Cidade
Universitária
05508-090 São Paulo - SP -
Brazil
55 11 3038-8571
broinizi@ime.usp.br

João Eduardo Ferreira
Department of Computer
Science
University of São Paulo
Rua do Matão, 1010 - Cidade
Universitária
05508-090 São Paulo - SP -
Brazil
55 11 3091-6111
jef@ime.usp.br

Alfredo Goldman
Department of Computer
Science
University of São Paulo
Rua do Matão, 1010 - Cidade
Universitária
05508-090 São Paulo - SP -
Brazil
55 11 3091-6111
gold@ime.usp.br

ABSTRACT

The creation of conceptual data design that appropriately represents specific application domain is one of the main challenges in requirements engineering. An initiative to help designers is the *Naked Objects* framework, where it is possible to interact with conceptual model in a limited way. The interactions are restricted to entity creations and single object-relations. We created an extension of the Naked Objects framework using annotations to allow manipulation of higher level abstractions as specialization and object-relationship. These abstractions allow better interactions between the domain specialist and designers. The use of our approach to explore and validate data requirements has several benefits: 1) It reduces conceptual specification problems (like poorly data requirements identification); 2) It narrows the distance among domain and design specialists; 3) It allows the simultaneous exploration of the conceptual data design and the system requirements.

1. INTRODUCTION

The creation of conceptual data design involves transformation of real problems into representations that can be implemented [6]. The process abstracts data from the real world in order to build a schema to represent the data. Semantically integrated data abstractions compose the schema to represent the real data [10].

We base our concepts on Naked Objects approach which reduces requirements identification problems and improves conceptual data design precision and agility. In order to reach these goals, the domain specialist participation is essential to validate the conceptual data design.

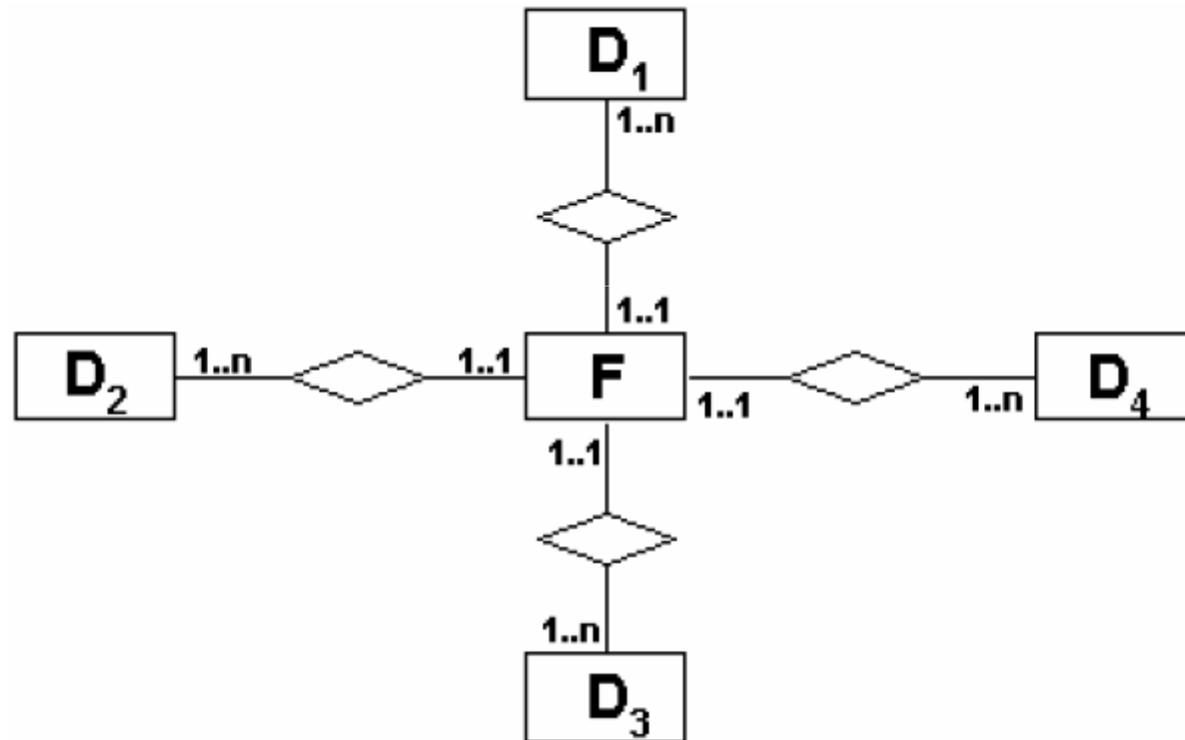
Our validation proposal is based on data abstractions [10] in order to improve precision of the conceptual data design. Besides precision, the agility of validation process is also improved by the use of concepts inspired on the *Naked Objects* (NO) [15]: *the role of the user of a system should be as a problem-solver*. The domain specialist is the only one





Fronteira

Modelos de Dados Multidimensional





Fronteira

Modelos de Dados Multidimensional

Definição 2.3.1 *Uma base de dados multidimensional é uma coleção de relações D_1, \dots, D_n, F , onde:*

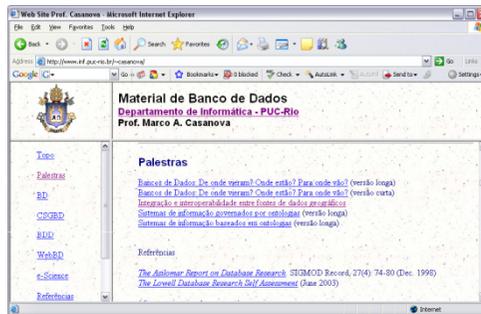
- *Cada D_i é uma tabela dimensão, isto é, uma relação caracterizada por um identificador que identifica unicamente cada tupla (d_i é a chave primária de D_i).*
- *F é uma tabela fato, isto é, uma relação que conecta todas as tabelas D_1, \dots, D_n ; o identificador de F é composto pelas chaves estrangeiras d_1, \dots, d_n de todas as tabelas dimensão conectadas. O esquema de F contém um conjunto de atributos adicionais V (que representam os valores sobre os quais serão aplicadas as funções de agregação).*





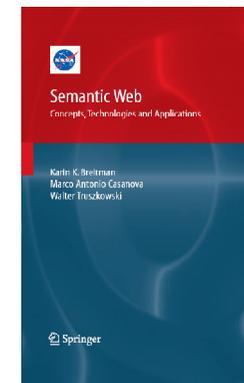
Fronteira

Cooperativas de Fontes de Dados



Casanova, M.A. *Revisitando Mapeamento entre Esquemas.*
<http://www.inf.puc-rio.br/~casanova/>

Breitman, K.; Casanova, M.A.; Truskowski, W.
Semantic Web: Concepts, Technologies and Applications.
Springer (2007)



Casanova, M.A.; Breitman, K.K.; Brauner, D.F.; Marins, A.L.A.
Database Conceptual Schema Matching.
IEEE Computer (Oct. 2007)



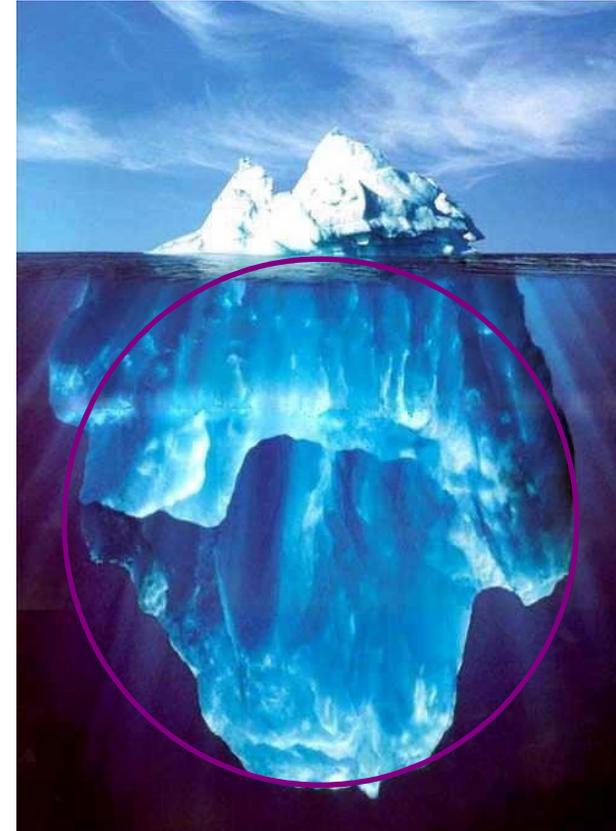
IME - USP



Fronteira

Cooperativas de Fontes de Dados

- Cooperativa de Fontes de Dados – CFD
 - coleção integrada de serviços sobre um conjunto de fontes de dados
- Desafios
 - 1) oferecer consultas, de forma simples, aos dados disponíveis
 - 2) expandir os serviços além de consultas
- Condições de contorno
 - 1) manter independência das fontes de dados
 - 2) permitir inclusão de novas fontes de dados
 - 3) acomodar um grande número de fontes de dados



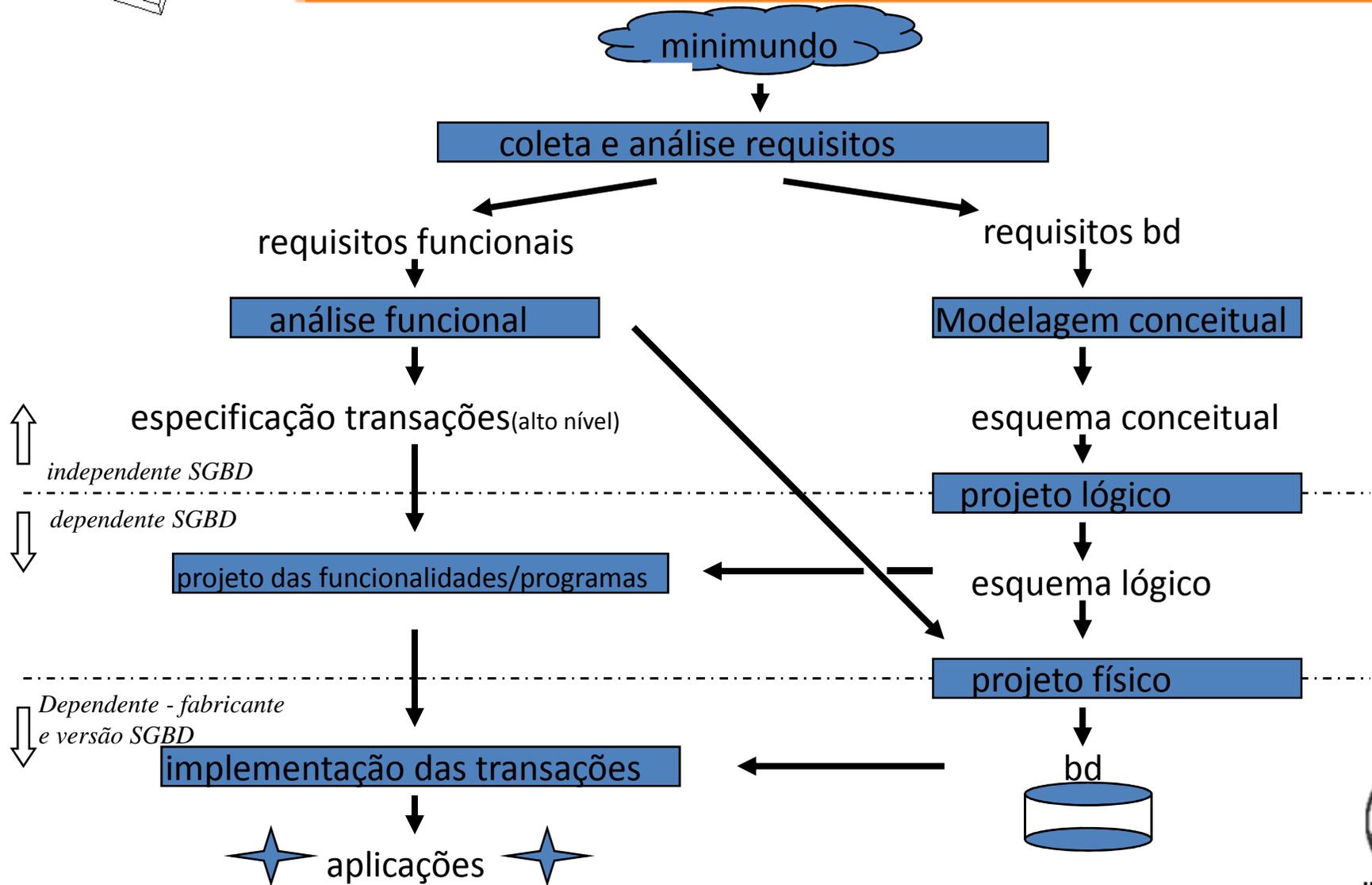
Iceberg Photo

Judith Currelly, Diane Farris Gallery





Próxima Aula





Próxima Aula

- Elmasri, R.; Navathe, S. B. [Trad.]. Sistemas de bancos de dados. Traduzido do original: FUNDAMENTALS OF DATABASE SYSTEMS. São Paulo: Pearson(Addison Wesley), 2005. 724 p. ISBN: 85-88639-17-3. Parte 2, páginas 89 – 205.
- Heuser, C.A., Projeto de Banco de Dados., Sagra - Luzzatto, 1 edição, 2004. Páginas 47 – 85.
- Ferreira, J.E.; Finger, M., Controle de concorrência e distribuição de dados: a teoria clássica, suas limitações e extensões modernas, Coleção de textos especialmente preparada para a Escola de Computação, 12^a, São Paulo, 2000. Cópia em: <http://www.ime.usp.br/~jef/ec2000.ps>
Cap 8. Páginas 82 – 110.



Fim



IME - USP